

## **AVALIAÇÃO DO CRESCIMENTO DE PEIXES SEM DADOS DE IDADE: UMA PERSPECTIVA BAYESIANA**

Renan Douglas Floriano Scavazzini (PIBIC/CNPq/FA/Uem), Diego Corrêa Alves (Orientador), Brian Alvarez Ribeiro de Melo (Coorientador), e-mail: renanscavazzini@gmail.com, dcalves@uem.br, barmelo2@uem.br.

Universidade Estadual de Maringá / Centro de Ciências Exatas/Maringá, PR.

**Probabilidade e Estatística / Probabilidade e Estatística Aplicadas.**

**Palavras-chave:** Biometria, Ecologia populacional, Ciência pesqueira.

### **Resumo**

No Brasil, o mercado de peixes produz R\$ 14,64 bilhões por ano, sendo esperado, no mundo, que este mercado atinja R\$ 1,121 trilhão em 2026. Uma dificuldade que restringe a gestão dos recursos pesqueiros é a determinação da estrutura etária dos estoques pesqueiros, uma vez que a obtenção desta é um processo oneroso e muitas vezes impreciso. Modelos estatísticos de crescimento somático são capazes de estimar não só a estrutura etária dos peixes como também o crescimento somático de um estoque, outro fator importante no manejo pesqueiro. Os modelos encontrados na literatura apresentam problemas quanto à suposição da distribuição do comprimento dos peixes e o tamanho amostral. Neste sentido, o presente trabalho busca expandir esses modelos generalizando para diferentes distribuições de probabilidade e realizando a estimação através de técnicas de inferência Bayesiana, afim de solucionar esses problemas.

### **Introdução**

O mercado mundial de peixes movimentou R\$ 796,88 bilhões em 2018, sendo esperado atingir R\$ 1,121 trilhão em 2026. É estimado que mais de 5 milhões de brasileiros estejam envolvido neste setor, gerando mais de R\$ 14,64 bilhões de produtos e serviços no país por ano. Contudo, a gestão dos recursos pesqueiros brasileiros ainda é precária, a qual deveria ser baseada em evidências a partir de estudos de dinâmicas populacionais e avaliações destes estoques pesqueiros. A estrutura etária de uma população é um dos principais componentes de entrada destes, possibilitando estimar recrutamentos e mortalidade. Entretanto, obter os dados de idade dos peixes é trabalhoso, caro e pode ser impreciso. Como alternativa foram desenvolvidos métodos estatísticos que permitem estimar estes parâmetros, sem nenhum dado de idade, apenas com dados de distribuição de tamanhos dos indivíduos, que são de obtenção mais rápida, fácil e barata. Alves et al.

(2018) propuseram um modelo estatístico baseado na suposição de normalidade da distribuição das frequências de comprimento dos peixes, o qual permite estimar tanto a estrutura etária quanto os parâmetros de crescimento somático a partir de uma única amostra de frequência de comprimento, e realizaram a estimação através de um método de estimação frequentista.

Embora a distribuição Normal tenha propriedades importantes, pode ser problemática em algumas situações como, por exemplo, permitir valores negativos em situações indesejadas e supor simetria, que pode ser delicado, especialmente com o tamanho amostral reduzido. Baseado nisso, buscou-se nesse trabalho generalizar os modelos de crescimento somático de peixes, propostos por Alves et al. (2018), para diferentes distribuições de probabilidade, tanto discretas quanto contínuas e realizar a estimação por meio de técnicas de inferência Bayesiana. Foram consideradas como candidatas a modelar a frequência de comprimento dos peixes as distribuições Poisson e Binomial Negativa, além da distribuição Normal, mas considerando duas diferentes estruturas de variância: homogênea e heterogênea.

## Materiais e métodos

Na modelagem foi utilizada a equação de crescimento de Von Bertalanffy (VBGE), cuja parametrização é dada por  $L_{\infty}(1 - e^{-K(t-t_0)})$ , onde  $L_{\infty}$  é o comprimento máximo esperado,  $K$  é o parâmetro de curvatura;  $t$  é a idade do peixe e  $t_0$  é a idade teórica quando o comprimento é igual à zero. O comprimento do peixe ( $s$ ) é uma variável aleatória, dependente da idade ( $t$ ), com a suposição de normalidade,  $s|t \sim N(L_{\infty}(1 - e^{-K(t-t_0)}), \sigma_{VBGE}^2)$ , onde  $\sigma_{VBGE}^2$  é a variância de  $s$  e foi considerado fixo e conhecido, assim como o parâmetro  $t_0$ , conforme proposto por Alves et al., (2018). Para construir o modelo, os comprimentos e as idades foram agrupados em intervalos de classes, sendo  $L$  um vetor coluna de tamanho  $m$  representando a classe de comprimento (em centímetros) e  $A$  um vetor coluna de tamanho  $r$  representando a classe etária (em anos). A matriz de transição é uma função da VBGE de tal forma que o elemento da linha  $i$  (comprimento) e coluna  $j$  (idade) serão dados pela integral da densidade de  $s|t$  em relação à  $s$ , no intervalo de  $L_i$  até  $L_{i+1}$ ,  $i = 1, \dots, (m - 1)$ . Com isso, a composição etária poderá ser estimada através da teoria dos modelos lineares ao definir o valor esperado de  $L$  como uma função da estrutura etária  $A$  e da matriz de transição  $P$ , ou seja,  $E(L) = PA$ .

Sob a perspectiva bayesiana, é necessária a especificação de distribuições *a priori* para os parâmetros que serão estimados. Foram assumidas distribuições Normais truncada (para que não haja valores negativos) e vagas (com elevada variância), sendo assim, distribuições *a priori* pouco-informativas:  $K \sim N(0,5; 0,0001)$ , truncada entre 0 e 5,  $L_{\infty} \sim N(120; 0,0001)$ , truncada no intervalo 0 a 250. Para a distribuição etária ( $A$ ) foi assumido que  $A \sim Dirichlet(\alpha_1 = 1, \alpha_2 = 1, \dots, \alpha_r = 1)$ . Já para os parâmetros de dispersão

dos modelos foram atribuídas distribuições *a priori*  $\tau_L \sim Gama(0,001; 0,001)$  e  $q \sim U(0,2; 4)$  para os modelos Normal homogêneo e heterogêneo, respectivamente. Para o modelo Poisson, a variância é igual a média, não contendo parâmetro extra de dispersão, enquanto que para o parâmetro de dispersão do Binomial Negativo foi assumido  $r \sim Gama(0,001; 0,001)$ .

A eficácia dos modelos foi avaliada com base em 100 conjuntos de dados, contendo 150 classes de comprimento e 9 classes etárias diferentes em cada conjunto, os quais são oriundos de populações virtuais da espécie de peixe *Brachyplatystoma rousseauxii*, conforme utilizado por Alves et al. (2018), onde os parâmetros da VBGE são dados por  $L_\infty = 153,3$ ,  $K = 0,22$  e  $t_0 = -0,49$ . Para o estudo de simulação foi utilizado a linguagem de programação *Python* na versão 3.8 e o *software* JAGS, um programa que faz simulações bayesianas com o algoritmo de amostrador de Gibbs. Para cada um dos modelos foram especificadas 100000 iterações, 4 cadeias de Markov executadas em paralelo, uma quebra de tamanho 100 e 5000 amostras de aquecimento. Foi utilizado um processo de multitarefa para rodar dois conjuntos de dados de uma só vez e um computador de processador Ryzen 7 1700 de 8 núcleos, 16 threads, 16 gb de memória ram e sistema operacional Linux.

Como critério de convergência dos modelos foi utilizada apenas a estatística  $R^2$ , assumindo que convergiram os ajustes que apresentaram  $R^2 \leq 1,15$  para os parâmetros  $L_\infty$  e  $K$ . Para comparar a qualidade do ajuste entre os modelos, foi utilizado o critério de informação de Watanabe (WAIC). O indicador utilizado foi a proporção de vezes que cada modelo apresentou o maior valor de WAIC entre os quatro modelos quando ajustaram ao mesmo vetor de dados. A performance dos métodos de estimação foi avaliada considerando o viés e tanto para o cálculo do WAIC quanto para o viés serão considerados, para cada modelo, apenas os conjuntos que apresentaram convergência.

## Resultados e Discussão

Na Tabela 1 observa-se que o tempo demandado na simulação com 100 repetições de cada modelo não variaram muito, demorando entre 6,7 e 8,8 horas. Os modelos que mais demoram foram o Normal Heterogêneo e o Binomial Negativo, sendo esses os mesmo que tiveram a maior proporção de convergências. Contudo, o modelo Normal Heterogêneo obteve um melhor ajuste na maioria das vezes. Na Tabela 2 nota-se que os vieses dos parâmetros foram menores nos modelos Normal de variância heterogênea e Binomial Negativo, sendo assim, obtiveram a melhor performance.

**Tabela 1** – Tempo total de execução, proporção de modelos que convergiram e proporção de vezes que o modelo apresentou o melhor ajuste (maior WAIC).

Modelo	Horas	Convergência	WAIC
--------	-------	--------------	------

Normal Homogêneo	6,7	14%	0,0%
Normal Heterogêneo	8,5	100%	87,0%
Poisson	6,8	31%	0,0%
Binomial Negativo	8,8	87%	11,5%

**Tabela 2** – Valor do viés entre os modelos que convergiram.

Parâmetro	Normal Homogêneo	Normal Heterogêneo	Poisson	Binomial Negativo
$L_{\infty}$	7,3643	1,0085	4,1162	0,9373
$K$	0,1002	0,0047	0,0772	0,0040
Idade (0,1]	0,5822	0,0003	0,4116	0,0007
Idade (1,2]	0,3490	0,0002	0,2680	0,0014
Idade (2,3]	0,1059	0,0002	0,0702	0,0006
Idade (3,4]	0,2527	0,0002	0,1637	0,0008
Idade (4,5]	0,0061	0,0003	0,0019	0,0003
Idade (5,6]	0,1002	0,0003	0,0684	0,0003
Idade (6,7]	0,0073	0,0003	0,0005	0,0003
Idade (7,8]	0,0218	0,0006	0,0160	0,0003
Idade (8,9]	0,0004	0,0001	0,0001	0,0001

## Conclusões

A proporção de modelos que convergiram é maior entre os que melhor acomodam melhor as variâncias, sendo eles o modelo Normal Heterogêneo e o modelo Binomial Negativo, além disso, a performance desses dois modelos é melhor do que os demais. Contudo, o modelo Normal Heterogêneo obteve um melhor ajuste em 87% das vezes, enquanto que o Binomial Negativo apenas 11,5% das vezes. Baseado nisso, podemos concluir que a distribuição Normal Heterogênea foi a mais adequada entre as distribuições candidatas para modelar e avaliar o crescimento de peixes sem dados de idade no problema simulado. Em uma próxima etapa deste estudo pretende-se testar outros algoritmos de estimação, com o intuito de obter ajustes melhores e mais rápidos, além de realizar um novo estudo de simulação avaliando o impacto da especificação de priores mais informativas, e, utilizando um processo de amostragens viesadas da população virtual.

## Agradecimentos

Os autores agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), à Fundação Araucária (FA) e à Universidade Estadual de Maringá (UEM) pelo suporte financeiro ao projeto e ao professor Dr. Job Diógenes Ribeiro Borges, filiado ao Nupélia UEM, pelo auxílio na implementação e execução dos códigos.

## Referências



Alves, D. C., Vasconcelos, L. P., Agostinho, A. A. Age composition and growth without age data: a likelihood-based model. **Fisheries Research**, 2014: 361-370, 2018.